Application for

Reissue of

United States Letters Patent

6,470,432 B2

S P E C I F I C A T I O N

EXTERNAL STORAGE CONTROL DEVICE AND DATA TRANSFER METHOD
BETWEEN EXTERNAL STORAGE CONTROL DEVICES

Koji OZAWA, of Odawara, Japan;
Kazuhide SANO, of Odawara, Japan;
Takeshi KOIDE, of Odawara, Japan; and
Katsunori NAKAMURA, of Odawara, Japan.

US 6,470,432 B2

1

EXTERNAL STORAGE CONTROL DEVICE AND DATA TRANSFER METHOD BETWEEN EXTERNAL STORAGE CONTROL DEVICES

This is a continuation application of U.S. Ser. No. 09/375,357 filed Aug. 17, 1999, now U.S. Pat. No. 6,321, 292, which is a divisional application of U.S. Ser. No. 08/779,471 filed Jan. 8, 1997, now U.S. Pat. No. 5,978,890 issued Nov. 2, 1999. The entire contents of these prior applications are hereby incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to an external storage control device which stores data to an external storage device according to a write command from a host. In particular, the present invention relates to a system in which the same data is copied to another external storage device.

BACKGROUND OF THE INVENTION

When data used in a host are stored to plural external storage devices, main and sub external storage devices that hold the same data are sometimes provided under individual external storage control devices. In this case, these external storage control devices are mutually connected, and the main external storage control device issues a write command to the external storage control device which controls sub external storage devices when the external storage control device which controls main external storage devices receives a write command. Thus, data stored in the main and sub external storage devices are duplicated.

U.S. Pat. No. 5,155,845 discloses a method in which an external storage control device which controls main external storage devices and one which controls sub external storage devices are mutually connected. A main external storage control device which has received a write command from a host, transfers data to a sub external storage control device. Thus, the write process is performed in parallel for both of main and sub external storage devices.

SUMMARY OF THE INVENTION

When a host handles data stored in external storage devices with CKD (count, key, data) format as used in large scale computer systems, the host issues channel commands in succession for instructing data transfer of each individual record. Thus, each individual record undergoes the same command chaining sequence in order to be transferred.

FIG. 3 illustrates a case of plural records being written using the same command chaining sequence. This Figure illustrates the case in which plural records are written in succession by a single command-chaining. In an external storage sub-system having main and sub external storage devices that hold the same data, located under individual external storage control devices, when a channel device that is a host issues a command-chain (DEFINE EXTENT/ LOCATE RECORD/WRITE(R1)/WRITE(R2)/WRITE (R3)) to write 3 successive records R1, R2, and R3 to a disk storage device that is under a main external storage control device, data flow between the channel device and the main external storage control device and between the main external storage control device and the sub external storage control device are shown in the processing sequences of FIG. 3. Thus, several command-chains between the main and sub external storage control devices for each data transfer of write records is executed.

Further, in case data is duplicated by adopting the system described above, the distance between main and sub external

2

storage control devices becomes large considering the back-ups necessary in case of a disaster. Thus, an optical fiber cable is adopted as the interface cable. Accordingly, the influence of cable delay, which is considered to be constant for a metal cable, with respect to command-chaining time cannot be ignored. However, since this duplication is for back up purposes, the influence of the write process for the sub external storage control device during an ordinary process must be minimized.

In a case where the write command must be executed a number of times in accordance with the command-chaining described above, this command-chaining time required for writing to sub external storage control devices cannot be ignored. This is because the amount of command-chaining between main and sub external storage control devices increases. As a result, backup processing can severely decrease the throughput rate in ordinary processing.

Thus, one purpose of the present invention is to optimize the write process time for the sub external storage control device by taking into consideration the command-chaining time between main and sub external storage control devices. As a result, the present invention offers a means for achieving excellent performance even under the conditions explained above.

In order to achieve the above-mentioned purpose, the external storage control device according to the present invention is equipped with a means for estimating command-chaining time between main and sub external storage control devices. There is also provided a means for estimating the time for a write process to a sub external storage control device before starting the write process to the sub external storage control device. The command is issued to the main external storage control device from the host. The present invention also includes means to select the best suited command-chaining method. Thus, a comparison is made between the above-mentioned estimated time with the data processing time required in the case of transferring data for a write record or data of plural tracks including a write record in one operation using a specified command. The present invention also includes command means for writing said data in a single operation.

It is possible for an external storage device to learn of the command-chaining time mentioned above by either measuring the command-chaining time from a specified command to the next command, or by setting the length of interface cable between main and sub external storage control devices from outside in advance. On the other hand, the time required for a write process for the sub external storage control device can be calculated by the above-mentioned information and by the information included in the command issued to the main external storage control device. Namely, command-chaining used for data input/output includes at least two specific channel commands prior to the command to start data transfer. For example, for an external storage control device, the commands, "DEFINE EXTENT" and "LOCATE RECORD", are issued prior to data transfer, and the number of records to be processed and the data length are given, so that the amount of data to be transferred can be calculated.

Furthermore, the command-chaining time between main and sub external storage control devices depends upon length of interface cable and the performance of the external storage control device, and does not depend on commands made before and after the command-chaining. Accordingly, command-chaining time during data transfer can be estimated by measuring the time for two command chains. Also,

3

the length of interface cable between main and sub external storage control devices, and support functions of a given external storage control device are known at the time of installation of a backup system. Thus, command-chaining time can be estimated with the length of interface cable as established.

Since the command-chaining time can be estimated as mentioned above, the processing time can also be estimated for a given amount of data transfer in the case that the command-chaining command from the channel device is issued to the sub external storage control device. Consequently, by comparing this processing time with a processing time in the case of transferring data for a write record, or, data of plural tracks including a write record together, using a command and command means to write said data at once, the best suited command-chaining instruction can be issued to the sub external storage control device.

Also, the external storage control device according to the present invention is equipped with means to confirm that the object data exists in the data buffer of another external storage control device in order to transfer data from the data buffer within the external storage control device to an external storage control device that receives the write commands, in addition to data transfer means as mentioned above. Moreover, the processing time can be further reduced by transferring all the physical data together, including the control byte and the check byte in the data buffer.

Additionally, in the case where object data for the read command, received from the host, exists in the buffer of an external storage control device other than the external storage control device which has received the read command, the processing time for performing a read-out can be shortened by transferring these data to the external storage control device which received the read command instead of accessing the other external storage device.

These and other objects, features and advantages of the present invention will become more apparent in view of the following detailed description of the preferred embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 illustrates a configuration of the data processing system according to an embodiment of the present invention.

FIG. 2 illustrates a configuration of the main disk control device according to an embodiment of the present invention.

FIG. 3 illustrates the data flow between the channel device, that is a host, the main disk control device, and the sub disk control device.

FIG. 4 illustrates the data flow in the case that data are transferred between the main and sub disk control devices record by record or in an entire track.

FIG. 5 illustrates data format in the data buffer of a disk control device for the case of 8 records per track.

FIG. 6 illustrates a graph showing the relation between number of records transferred and the processing time.

FIG. 7 illustrates a flow chart representing steps followed according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The preferred embodiments of the present invention will now be described in conjunction with the Figures. FIG. 1 illustrates a configuration diagram of a data processing

4

system according to an embodiment of the present invention. This data processing system includes a main memory device 1, a channel device 2, a backup system 6 and a channel device 7. This data processing system also includes a main disk control device 3 equipped with a buffer memory 4, and a main disk storage device 5, and also includes a sub disk control device 8 equipped with a buffer memory 9, and a sub disk storage device 4. The channel device 2 is connected to the main disk storage control device 3 through an interface cable 11, the main disk storage control device 3 is connected to the sub disk storage control device 8 through an interface cable 12. Interface cables are also used to interconnect other elements shown in this Figure. Preferably, optical fiber cables are used as the interface cables.

As shown in FIG. 2 the main disk control device 3 includes a channel command analyzing section 31, a command judging section for judging commands to be issued to the sub disk control device 32, a management information table for the sub disk control device 33, a timer 34, a command issuing section for the sub disk control device 35, and a section for interpreting cable length and the function of the sub disk control device 36. Function blocks 31 33, 35 and 36 are realized by a micro-program executed by a microprocessor contained in the disk control device 3. Timer 34 is a hardware counter that is counted up according to a constant period. The sub disk control device 8 is constructed similarly to main disk control device 3.

The command-chaining time required between main and sub external storage control devices is measured in the main disk control device 3, by using the timer 34, the section for issuing command to the sub disk control device 35, and the section for interpreting cable length and function of the sub disk control device 36. Then, the command chaining time, Tsg, is measured when a command is processed once a pair of main and sub disk storage devices are established. The measuring is performed by timer 34 and is stored in the management information table 33 in the main disk control device 3.

The main disk control device obtains support function level information of the sub disk control device when the pair of main and sub disk storage devices is established. This function level information is stored in the management information table of the main disk control device. This stored information is used to judge whether a write command to write data for write record, or, data of plural tracks including the write record can be accepted by the sub disk storage control device or not.

If the command judgment section for the sub disk control device 32 has judged from the support function level information that a specified command is not acceptable, when a write command is received from the channel device 2, the section for issuing command to the sub disk control device 35 unconditionally issues a command-chain sent from the channel device to the sub disk control device. Thus, the system is rendered more versatile in that it can connect even with a conventional device that does not support these functions.

The channel command analyzing section 2 operates as follows. When the channel device 2 issues a write command to the main disk control device 3, the channel command analyzing section 31 determines whether or not the write command is for a disk storage device that forms a pair with the main disk control device 3.

FIG. 6 is a graph illustrating command-chaining time for two separate cases. One case corresponds to the processing time required for transfer of data for one track including a